# Graphic Analysis of Crop Rotation Data

F. B. CADY and R. H. BROWN

BU-614-M June 1977

## Abstract

A preliminary analysis, emphasizing graphic techniques, used data from the comprehensive Aurora Research Farm Rotation Experiment (Hatch project 452). The analysis represented an initial attempt at a combined analysis of the yearly data. Two major objectives were attained. The data were punched from field sheets and edited. Plot data were calculated from the raw field results and are now available in a permanent and easily accessible form, either on cards or tape. Secondly, the data were subjected to various standard preliminary analyses. Emphasized in this report are techniques utilizing graphical procedures useful in Biometry and recently developed by Kronmal and Tarter. The preliminary analysis shows the major features of the data set and has been helpful in later work, including a more complete analysis by Baldock.

## 1. Introduction

A crop rotation is defined as a sequence of one or more crops on the same land. Most present day rotation experiments are designed to compare treatments applied to crops from one or more rotations. The motivation behind these experiments and the long-time interests of agriculturists is well expressed by the partial title of a recent dissertation by Baldock (1976), "What cropping system for New York dairy farms?" Baldock's work was based on the Aurora Research Farm Rotation Experiment carried out by Dr. Robert B. Musgrave and his colleagues during 1955-1968. The experiment compared various rotations ranging from continuous corn to continuous hay or silage. Small grains were also included. Most of the rotations were five years in length; consequently, at least two full cycles of each rotation after several preliminary years were carried out. In addition, fertilizer and manure treats had been applied. The rotation experiment is almost unique in that the same experimenters were involved during the entire period and other variable factors had been minimized. Annual interpretations of the data had been made but no analysis over the entire period had been attempted. The data from 1959-1968 are utilized and the first phase of the analysis is reported here. The Baldock dissertation utilizes the results from the first stage and develops a full analysis, including an economical evaluation.

The preliminary analysis, covered in this report, had two major objectives:

(1) Data management. Each year of the experiment resulted in 816 observations which were picked up from the original field sheets for the 14 years. After punching, the data were then edited and put on tape in a form usable by others in the future.

(2) Graphic analysis. With a large data set, the first phase of the data analysis was graphic, utilizing newly developed density plot methodology by Tarter and Kronmal (1976).

## 2. Terminology

Suppose a number of annual crops are grown in a pre-assigned order in successive years. The pre-assigned order may be repeated cyclically and may contain the same crop more than once. For example, a pre-assigned order may include the following crops appearing in that order: corn (C), corn and oats (O). If we denote the pre-assigned order by C-C-O, then C-C-O may be repeated cyclically as C-C-O-C-C-O-C ... .

Cycle Every repetition of the pre-assigned order constitutes a cycle, or a cropping cycle. In the example above, C-C-O is a cycle.

*Length* The length of a cycle is the number of years required for one repetition of the pre-assigned order of the cylce. The length of the cycle C-C-O is therefore three. For convenience, continuous cropping such as continuous corn is regarded as having a length of cycle of one year.

*Phase* Phase denotes the position of a crop in a pre-assigned order or a cycle. The number of phases of a cycle are equal to the length of the cycle. Thus there are three phases for the cycle C-C-O, the first and second phases being corn, and the third phase being oats.

*Sequence* If a number of crops are grown successively in cyclic order, then any portion of successive crops in that order form a sequence. For any given pre-assigned order, sequences beginning with the same phase are considered the same, and sequences beginning with different phases have a common cyclic order. Therefore, a pre-assigned order of n phases has n sequences, each headed by a different phase. For example, for the pre-assigned order C-C-O, there are three sequences C-C-O-..., C-O-C-... and O-C-C-..., denoted respectively by C-C-O, C-O-C and O-C-C.

*Rotation* A rotation is a management system of growing a number of crops in sequence on the same piece of land. The length of a rotation is equal to the length of the associated cycle. Sincere there are as many possible sequences as there are phases, one of the sequences is often chosen to represent the rotation and is referred to as the *basic rotation*. Hereafter in this paper, we shall often use the term rotation to describe a specific phase of a basic rotation. With this definition the first- and second-year corn in the basic rotation C-C-O are different rotations and are denoted by $\underline{C}$-C-O and C-$\underline{C}$-O, respectively. The underlining notation is also useful in indicating the crop being considered in an analysis or a discussion. If the oats crop is being discussed, for example, the rotation may be denoted as C-C-O.

Test Crop   Since there are generally different crops in a basic rotation, it is convenient to refer to the crop under investigation as test crop or test. For example, either the corn or oats of the basic rotation C-O may be the test crop, depending on whether the corn or oats is under investigation.

Rotation Effect and its Measurement   Consider a fertility experiment. The effects on crop yield of a fertility treatment applied in the current year are known as direct effects, while the effects on the yield of current crop of treatments applied in preceding years are known as residual effects. The combined effects of treatments applied currently and in previous years are called cumulative effects. If the same fertility treatment is applied to the same land in successive years, the cumulative effects will increase over the years at a decreasing rate. Eventually, after a theoretical equilibrium between the fertility treatment and soil fertility is attained, the cumulative effects will remain constant.

A rotation system affects the soil fertility in much the same way except that instead of years, the direct, residual and cumulative effects are defined in terms of cycles. In particular, cumulative effects will be referred to as rotation effects. As the objective of the adoption in practical situations of a rotation system is to increase or maximize long-term farming profit, the rotation effects have an upward trend and will stabilize in the long run. There are rotation systems, however, that have negative direct and residual effects. These systems cause deterioration of soil fertility over the cycles at a decreasing rate, and therefore the rotation effects have a declining trend. Whether rotation effects increase or decrease as the cycles advance, the trend can be, and has long been, measured in terms of the change in crop yield. Cochran (1939) has demonstrated the use of linear and quadratic terms to represent the trend of yield changes. A similar approach was also used by Patterson (1964).

As rotation effects stabilize, the yield of the rotation approaches a constant, called the limiting or asymptotic yield. Associated with any rotation is a limiting yield which, if known, will indicate the yield of the rotation one may expect in the long run. This is valuable information for a research worker wishing to compare the relative merits of various rotation systems or deciding when to close out a rotation experiment. An example is given by Fuller and Cady (1965).

## 3. Experimental Error Structure

In statistical analysis of experimental data, a model is hypothesized or constructed for the phenomenon under investigation. The model generally consists of two parts, a mathematical function explaining the cause and response relationship between some input variables and an output variable, and an experimental error which accounts for the deviation of the observed cause and response relationship from the hypothesized mathematical function. The two parts are usually assumed to be additive, and may be expressed in the general form

$$y = f(X,\theta) + e \tag{1}$$

where y is the output or dependent variable, X is a matrix of input or independent variable, $\theta$ is a vector of parameters, $f(X,\theta)$ is a mathematical function to be specified in actual analysis, and e is the experimental error associated with y . While the main interest of a research worker may be to determine $f(X,\theta)$ and to estimate $\theta$, the importance of making assumptions for the error structure should not be overlooked because the applicable statistical procedure will depend on these assumptions. A commonly accepted assumption of experimental error is that the errors associated with different outputs are independent normal variables with zero mean and a common variance. The assumption is considered to be realistic for most annual experiments with rather simple designs and permits the use of statistical procedures based on the theory of least squares.

For understanding the error structure of rotation experiment models, consider a crop grown on p plots over w years. A reasonable model is to attribute part of the variability among the observations to years, part to plots and to assume that the residual, the year x plot interaction, is a random error, i.e.,

$$y_{ik} = q_k + p_i + w_{ik}$$

where $k = 1, 2, \cdots, w$ years and $i = 1, 2, \cdots, p$ plots. If the variability among plots is also assumed to be random error, then the experimental error, $e_{ik}$, is the sum of two components

$$e_{ik} = p_i + w_{ik} \quad .$$

The $p_i$ plot error is assumed to be constant over the years with a common variance, $\sigma_p^2$ . The $w_{ik}$ year x plot error is independent of years and plots with variance $\sigma_w^2$ . The variance of $e_{ik}$, $\sigma^2$, is then the sum of $\sigma_p^2$ and $\sigma_w^2$ . Zero covariance between $e_{ij}$ for different p plots is a reasonable assumption. Unfortunately the same zero covariance between $e_{ij}$ for different years on the same plot cannot be assumed. A characteristic of rotation experiments is that the same crop occurs on the same plot over a number of years, i.e., yields are repeated measurements made on the same experimental unit. Since the soil fertility of a plot in different years tends to be correlated, the yields observed on the same plot logically cannot be assumed to have independent errors. The covariance structure is then

$$\begin{aligned} \mathrm{cov}(e_{ik}, e_{i'k'}) &= \sigma^2 = \sigma_p^2 + \sigma_w^2 && \text{if } i = i',\ k = k' \\ &= \rho\sigma^2 = \sigma_p^2 && \text{if } i = i',\ k \neq k' \qquad (2) \\ &= 0 && \text{if } i \neq i' \ . \end{aligned}$$

Covariance structure (2) can be recognized as analogous to the error structure for a nested random effects model, i.e., $\sigma_p^2$ is the variance component due to the experimental units or the primary sampling units and $\sigma_w^2$ is the variance component due to samples within units. Note that the expected mean squares would be as usual except that the degrees of freedom for estimating $\sigma_w^2$ would be those associated with the plot X year interaction. Battese, Fuller and Shrader (1972), Björnsson and Cady (1973), and Patterson and Lowe (1970) give additional development.

When fertilization or management treatments are included in a rotation experiment, they generally appear as split plot treatments of a split plot design. Three components of variation are assumed to be associated with an observed yield. The plot errors, with variance $\sigma_p^2$, are components of the covariance between any two split plot yields observed on the same plot, regardless of whether or not they have the same split plot treatment. The year X plot errors, with variance $\sigma_w^2$, are defined as in (2). Thus only observations appearing on the same plot in the same year have $\sigma_w^2$ as a component of their covariance. A third source of variation is the split plot to split plot variation within the same plot. The variance of this variation is assumed to be $\sigma_s^2$ . Denoting the split plot treatments by the subscript j, the error assumption can be summarized as

$$\begin{aligned} \operatorname{Cov}(e_{ijk}, e_{i'j'k'}) &= \sigma_p^2 + \sigma_w^2 + \sigma_s^2 && \text{if } i = i',\ j = j',\ k = k' \\ &= \sigma_p^2 + \sigma_w^2 && \text{if } i = i',\ j \neq j',\ k = k' \\ &= \sigma_p^2 && \text{if } i = i',\ j \neq j,\ k \neq k' \\ &= 0 && \text{if } i \neq i' \ . \end{aligned} \tag{3}$$

where $j = 1, 2, \cdots, s$ . A correspondence between the variance components in (3) and the error structure of a nested model is that $\sigma_p^2$ is the variation due to experimental unit, $\sigma_w^2$ the variation due to sample, and $\sigma_s^2$ the variation due to determination.

With the inclusion of a split plot treatment a fourth component, $\sigma_t^2$, the split plot X year error, also has to be included. The error model is then assumed to be

$$e_{ijk} = p_i + s_{ij} + w_{ik} + t_{ijk} \tag{4}$$

where the errors $p_i$, $s_{ij}$, $w_{ik}$ + $t_{ijk}$ are independently distributed with zero means and variances $\sigma_p^2$, $\sigma_s^2$, $\sigma_w^2$ and $\sigma_t^2$ respectively. The covariance structure of $e_{ijk}$ is written as

$$\begin{aligned} Cov(e_{ijk}, e_{i'j'k'}) &= \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_t^2 && \text{if } i = i', j = j', k = k' \\ &= \sigma_p^2 + \sigma_s^2 && \text{if } i = i', j = j', k \neq k' \\ &= \sigma_p^2 + \sigma_w^2 && \text{if } i = i', j \neq j', k = k' \\ &= \sigma_p^2 && \text{if } i = i', j \neq j', k \neq k' \\ &= 0 && \text{if } i \neq i' \ . \end{aligned}$$

Without these covariance terms, the parameters of model (1) can be estimated by

$$\hat{\theta} = (X'X)^{-1}X'y \quad .$$

With rotation experiments, $E(ee')$ cannot be assumed to be $\sigma^2 I$ but is $\sigma^2 V$ where $V$ includes the covariance terms given previously. Generalized least squares estimation have to be used to obtain efficient, unbiased estimators. However, for the covariance structure assumed here, straightforward transformation of the data for removing the correlations among the errors are available. Simple least squares estimators are then used. This procedure is equivalent to generalized least squares but eliminates the need of inverting the matrix $V$ . Further development of the required transformations are given in Shih (1966), Battese and Fuller (1972), Battese, Fuller and Shrader (1972) and Fuller and Battese (1973).

Due to weather variables, results of agricultural field experiments show large year-to-year variability. In the analysis of rotation experiments, the usual procedure is to assign dummy variables for years (Cochran 1939, Patterson

1959 and Yates 1954). In one example Yates also partitions the year main effect into among series and years within series. Cady and Mason (1964) give further details on this type of analysis of variance approach.

The exponential response model approach used by Fuller and Cady (1965) estimated asymptotic or limiting yields and standard errors which are useful in practical applications. These Iowa rotation experiments gave support to an assumption of zero plot correlations leading to a simplication of the error structure. Yearly variation was handled directly through the estimated parameters and the error model was simply the sum of one variance component for the whole plot and a second component for the split plot. The required transformation was simpler than that required if model (4) had to be used. This approach was then used by Shrader, Fuller and Cady (1966) in testing a specific hypothesis and by Battese, Fuller and Shrader (1972) in an economic analysis. The results from the 1966 paper are particularly interesting since support was given to the conjecture that an estimated exponential function would fit data from various rotations, leading to estimates of the nitrogen contribution by the rotations.

## 4. Data Management

In its original format, data was punched from the field sheets to computer cards. The information which was taken from the field sheets was the year, the crop, the plot number and sub-plot identification, the plot size, the plot green weight and the sample green and dry weights. For the corn plots a stand count was also recorded. Green and dry weights for up to three cuts were punched for the hay yields and, when available, for straw. A computer program was written which, on the basis of the year, plot number and sub-plot identification, coded each observation as to rotation, phase of rotation, replication and treatment.

On the basis of the sample green and dry weights, the plot green weight was transformed into dry matter yield in kilos per hectare. This information, along with the actual levels of commercial and animal fertilizer applied, was added to the information already on the cards and new cards stored in the computer as card filed.

A card file is simply a set of data cards which the computer has read and stored semi-permanently on an external recording device called a disk. The data can then be used as input to any program without physically having the data cards in the deck setup. The user simply tells the computer to insert the card file that contains the data cards he wishes to use into the appropriate place in the job stream (i.e., wherever he would normally put the data cards themselves).

Each year of data was processed separately and is stored in separate card files. The card file names are four characters long and have the form YRab, where ab represents the year in which the data was collected. For example, the card file named YR59 contains all the data collected in 1959, the card file YR60 contains all of the data for 1960, and so on through YR68.

A detailed description of the format of the cards is given in Table 1. The original punched cards, containing only the information taken from the field sheets, are stored in Emerson 160. In order to maintain independent and permanent copies of all the data, the card files have been copied onto three magnetic tapes. Each tape contains all ten years of data. A detailed description of the way in which the tapes were written, their volume serial numbers, etc., is given in an appendix. The tapes are intended mainly as a form of permanent storage. For the purpose of input to a program, it is more cost-efficient to use the card files. The tapes are held by Jon Baldock, Foster Cady and Robert Musgrave and are available to others for their use.

Table 1. The Data

| Columns | Contents |
|---|---|
| 1-2 | year code (59-68) |
| 3-6 | sample green weight |
| 7-9 | plot size in square feet |
| 10-12 | alphabetic crop code (left justified) |

Alphabetic Crop Codes

| | |
|---|---|
| A - alfalfa | HR - hay ryegrass |
| CE - ear corn | HT - hay timothy |
| CS - corn for silage | OTG - oats for grain |
| HB - hay brome | OTH - oats for hay |
| HBF - hay birdsfoot | SBG - spring barley |
| HC - hay clover | WBG - winter barley |
| HCT - hay clover and timothy | WHG - wheat for grain |

| | | | |
|---|---|---|---|
| 16-18 | plot number code (100-633) | 21-25 | plot green weight, first cut hay |
| 19 | sub-plot code (N, O, R, or S) | 26-30 | sample dry weight, first cut hay |
| 20 | aspect code (blank, E, or W) | 31-35 | plot green weight, second cut hay |
| 21-25 | plot green weight, corn | 36-40 | sample dry weight, second cut hay |
| 26-30 | sample dry weight, corn | 41-45 | plot green weight, third cut hay |
| 31-35 | stand count, corn | 46-50 | sample dry weight, third cut hay |
| 21-25 | plot green weight, grain | 51-54 | level of N in lbs/acre |
| 26-30 | % dry matter, grain | 55-56 | level of P in lbs/acre |
| 31-35 | test weight, grain | 57-58 | level of K in lbs/acre |
| 36-40 | plot green weight, straw | 59-66 | dry matter yield in kilos/hectare |
| 41-45 | sample green weight, straw | 67-68 | year code (1-10) |
| 46-50 | sample dry weight, straw | 69 | blank |

| Columns | Contents |
|---|---|
| 70-71 | rotation code (1-13) |

| | | | | |
|---|---|---|---|---|
| A - 01 | $C_C$ - 04 | F - 07 | I(E) - 10 | J(W) - 13 |
| B - 02 | D - 05 | G - 08 | I(W) - 11 | |
| C - 03 | E - 06 | H - 09 | J(E) - 12 | |

| Columns | Contents |
|---|---|
| 72 | phase of the rotation coded 0 - 5. A phase of zero (0) means a continuous rotation. |
| 73 | manure coding. The amount of manure applied is two times the number in this column in tons/acre. |
| 74-75 | numeric crop code (1-13) |
| 76 | year of the crop within the rotation |
| 77 | blank |
| 78 | treatment numeric code (W - 1, X - 2, Y - 3, Z - 4) |
| 79 | replication code (1-6) |
| 80 | cycle code (1-3) |

## 5. Data Storage

Data for all ten years are currently maintained in two forms. One form is a set of ten card files, one for each year. The card files have the names YR59, YR60, ..., YR68, indicating which year of data is contained in that file. Any or all of ten card files can be used as input for analysis to a program by use of one or more HASP commands of the following form:

```
   /*INSERT ƀ QSQ.YR64   .
  ccl
```

Besides the card files, three separate copies of the data exist on three magnetic tapes. Each tape contains an entire copy of all ten years of data. The tapes are standard OS labeled and their volume serial numbers are QSQ01, QSQ02, and QSQ03. To use one of the tapes in a program one of the following HASP commands will be required:

```
 /*SETUP ƀ ID = QSQ01,  UNIT = TAPE9 ,
ccl
          TEXT = 'ROTATION ƀ A',  XL = QSQ01A

 /*SETUP ƀ ID = QSQ02 ƀ UNIT = TAPE9 ,
          TEXT = 'ROTATION ƀ DATA ƀ B',  KL = QSQ02B

 /*SETUP ƀ ID = QSQ03 ƀ UNIT = TAPE9 ,
          TEXT = 'ROTATION ƀ DATA ƀ C',  KL = QSQ03C  .
```

Each tape contains ten (10) files. The first file on each tape has the data set name YR59. The second file on each tape has the data set name YR60, and so on. Each tape is recorded at 6250BPI (DEN = 4) and each file is written with the following data control block information:

```
DCB = (RECFM = FB,  LRECL = 80,  BLKSEZE = 3200)  .
```

A detailed list of the file structure of each tape, including the number of logical records in each file, follows on the next page.

```
UORCOPY (V=07.06.70)    ALL DATA SET(S) TO BE COPIED

 INPUT: SYSUT1    DSN=YR59,LABEL=(1,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      1 COPIED,       840 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR60,LABEL=(2,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      2 COPIED,       840 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR61,LABEL=(3,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      3 COPIED,       816 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR62,LABEL=(4,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      4 COPIED,       840 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR63,LABEL=(5,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      5 COPIED,       864 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR64,LABEL=(6,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      6 COPIED,       864 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR65,LABEL=(7,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      7 COPIED,       870 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR66,LABEL=(8,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      8 COPIED,       866 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR67,LABEL=(9,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET      9 COPIED,       846 LOGICAL RECORD(S) PROCESSED

 INPUT: SYSUT1    DSN=YR68,LABEL=(10,SL),VOL=SER=QS001,UNIT=*<$/#@%
     DCB=(RECFM=FB,BLKSIZE=3200,LRECL=80)
OUTPUT: SYSUT2    DUMMY OUTPUT DATA SET ****
DATA SET     10 COPIED,       846 LOGICAL RECORD(S) PROCESSED
```

TAPE 9 (handwritten beside each of the ten DCB lines)

## 6. Graphic Analysis

### 6.1. Smoothed Histograms

The general purpose of a histogram is to discover how the distribution of a random variable, e.g., corn yields or total dry matter for a given year, is spread over a range of observed values. Graphical measures of central tendency and dispersion are obtained. Isolated or extreme data points can be identified. The histogram can also be considered as an estimate of the probability density function of the random variable. If more and more observations were taken and the histograms constructed by grouping into intervals of less and less width, we can imagine that the histograms would tend toward a smooth curve.

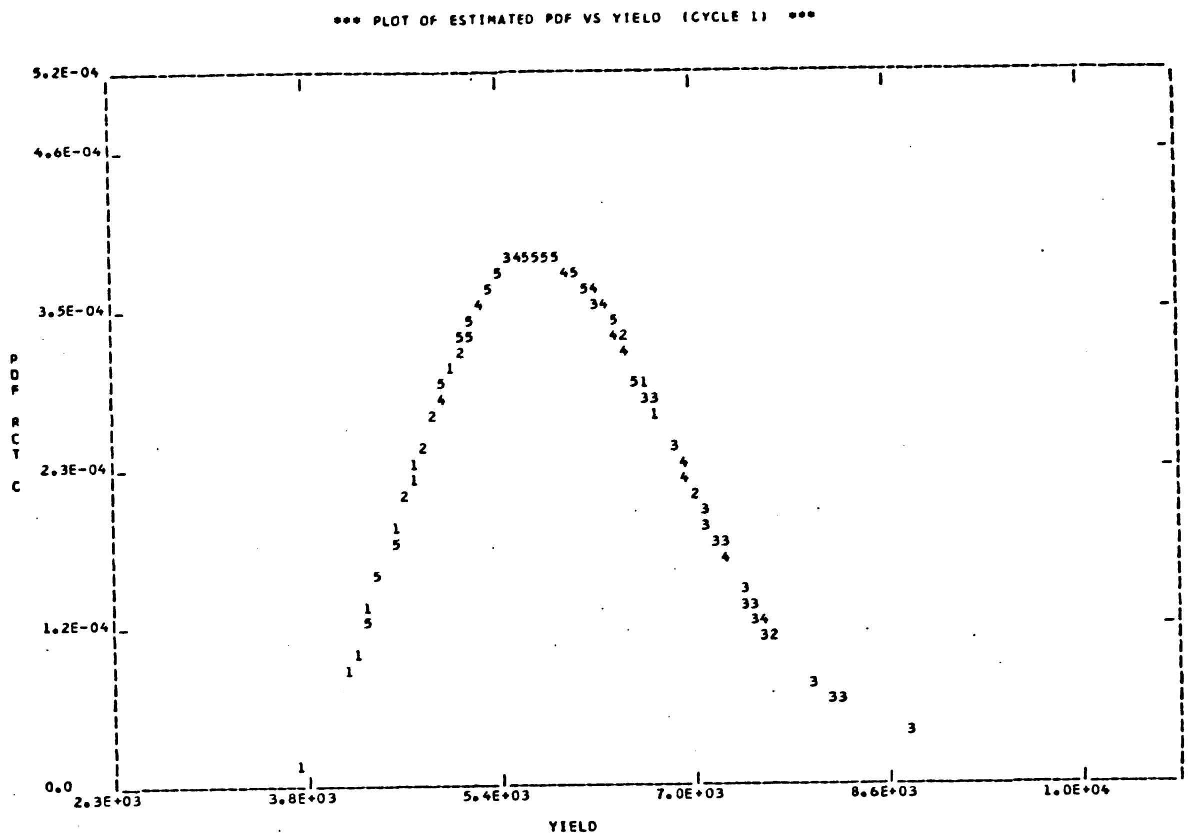
However, the non-smooth histogram based on a finite number of data points presents an excess of visual ones which seem to obscure its basic purpose. The viewer tends to spend more time mentally smoothing over the peaks and filling in the gaps than in trying to interpret the histogram itself. It is also difficult to assess other characteristics of a distribution (other than overall central tendency and dispersion) such as skewness or truncation.

It seems reasonable that if the histogram can serve as an estimate of the probability density function (pdf), then more general types of pdf estimates can serve the basic function of the histogram, i.e., to provide a graphical display of the distribution of a random variable. In fact the method used in constructing the printer plots of the corn yields is such an alternative type of density function estimate. It gives continuous and visually more appealing summaries of the corn yield distributions.

The method used is due to Kronmal and Tarter (1968) and is based on estimating coefficients in an orthogonal series expansion of an arbitrary continuous density function. The orthogonal series used is the trigonometric polynomial. Whereas, in the use of histograms the choice of interval lengths and location determines the fit of the graph to the data, in the present approach it is the degree of the approximating polynomial (i.e., the truncation point of the series expansion) which is arbitrary. The choice of the trigonometric polynomials leads to an estimate of the integrated mean squared error of the truncated estimate of the density function and hence to a criterion for choosing an optimal number of coefficients to be estimated. For a detailed description of the statistical basic of the technique and the numerical algorithm, see Kronmal and Tarter (1968) and Tarter, Holcomb and Kronmal (1967). An overall view is given in Tarter and Kronmal (1976).

The interpretation of the plots is essentially the same as the interpretation of a histogram. The numbers 1-5 are for years 1959-63 (cycle 1) and 6-9 and A are for years 1964-68 (cycle 2). Peaks are generally located around intervals in which there is a concentration (i.e., high frequency) of data points. Each plot potentially has 24 points (4 treatments and 6 replications). However, a printer usually cannot put two characters in the same position on the same line. When printer plots are generated and two observations do fall in the same position on the same line, that observation which is closer to the end of the data set will overwrite (or hide) the other observation on the printed line. Especially in the areas of the peaks on the pdf plots, the observations which are printed are hiding other observations. Here the impression of high (relative) frequency of occurrence is obtained from the peak itself rather than the actual frequencies. Notice (from the scale of the y-axis) that the heights of the peaks in the cycle two plots are considerably less than they are in the cycle one plots. Quantitatively fewer observations are going into these peaks since the good and bad years of cycle two are separating into their own distributions.

In looking at the cycle one density plots for the corn yields (Figures 1-3), some general conclusions can be made by comparing the two plots on a page or by placing one or more plots over each other and holding up to a light. The rotations are:


*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***
5.2E-04
4.6E-04
3.5E-04
2.3E-04
1.2E-04
0.0
P D F R C T C
2.3E+03
3.8E+03
5.4E+03
7.0E+03
8.6E+03
1.0E+04
YIELD

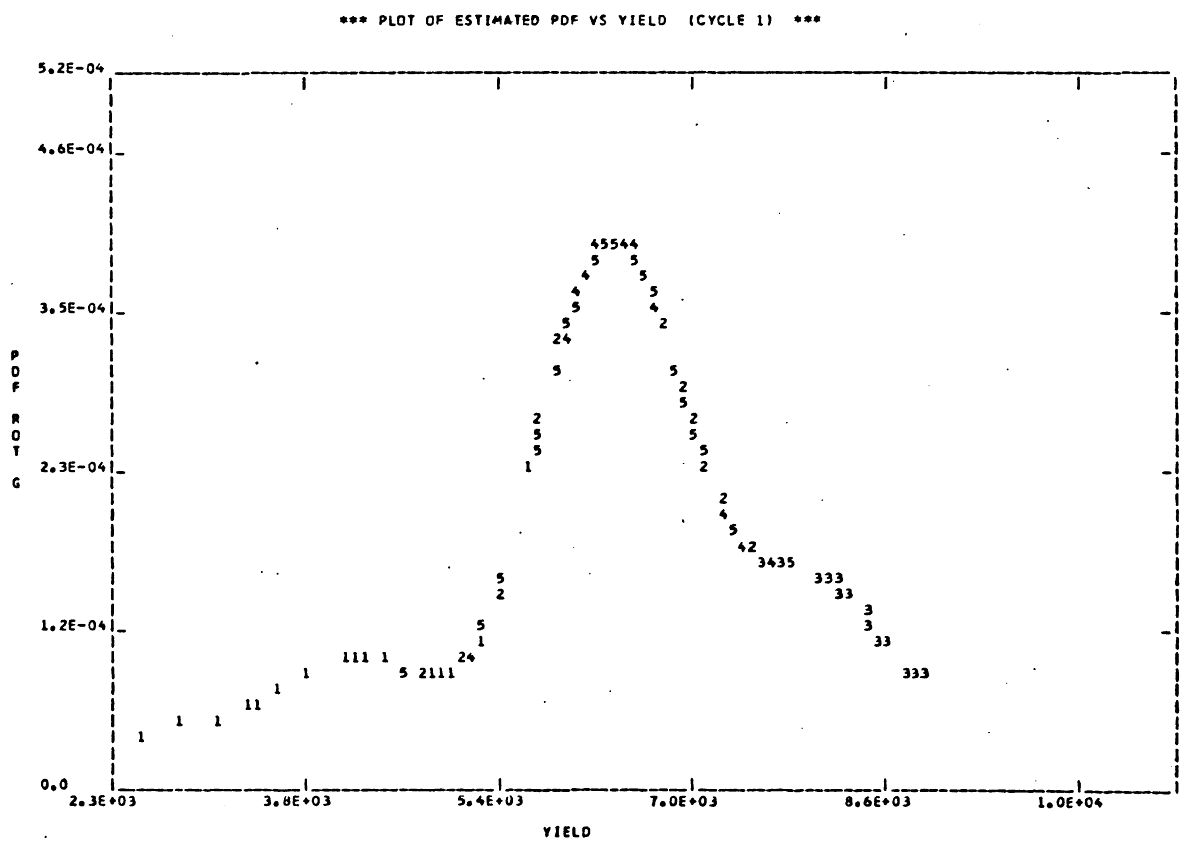
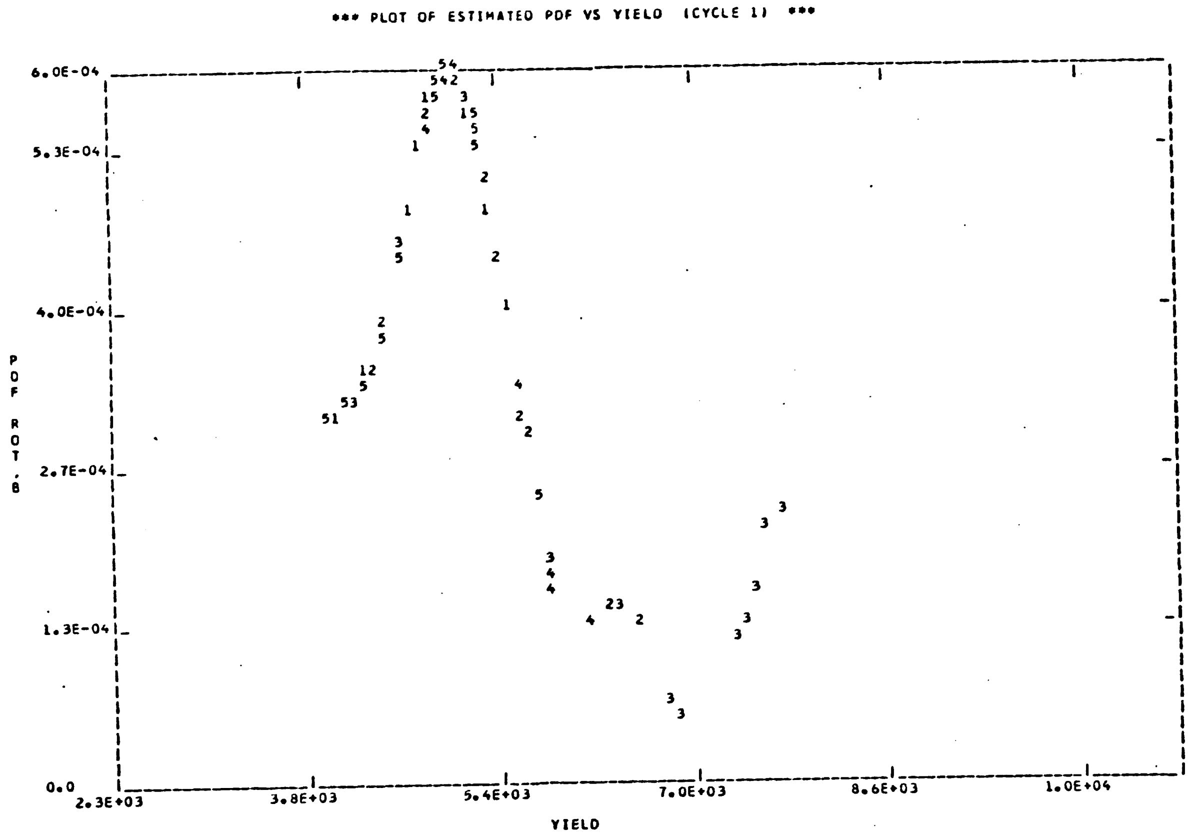

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***
5.2E-04
4.6E-04
3.5E-04
2.3E-04
1.2E-04
0.0
P D F R O T G
2.3E+03
3.8E+03
5.4E+03
7.0E+03
8.6E+03
1.0E+04
YIELD

Figure 1


*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***
PDF ROT. 8
YIELD
6.0E-04
5.3E-04
4.0E-04
2.7E-04
1.3E-04
0.0
2.3E+03
3.8E+03
5.4E+03
7.0E+03
8.6E+03
1.0E+04

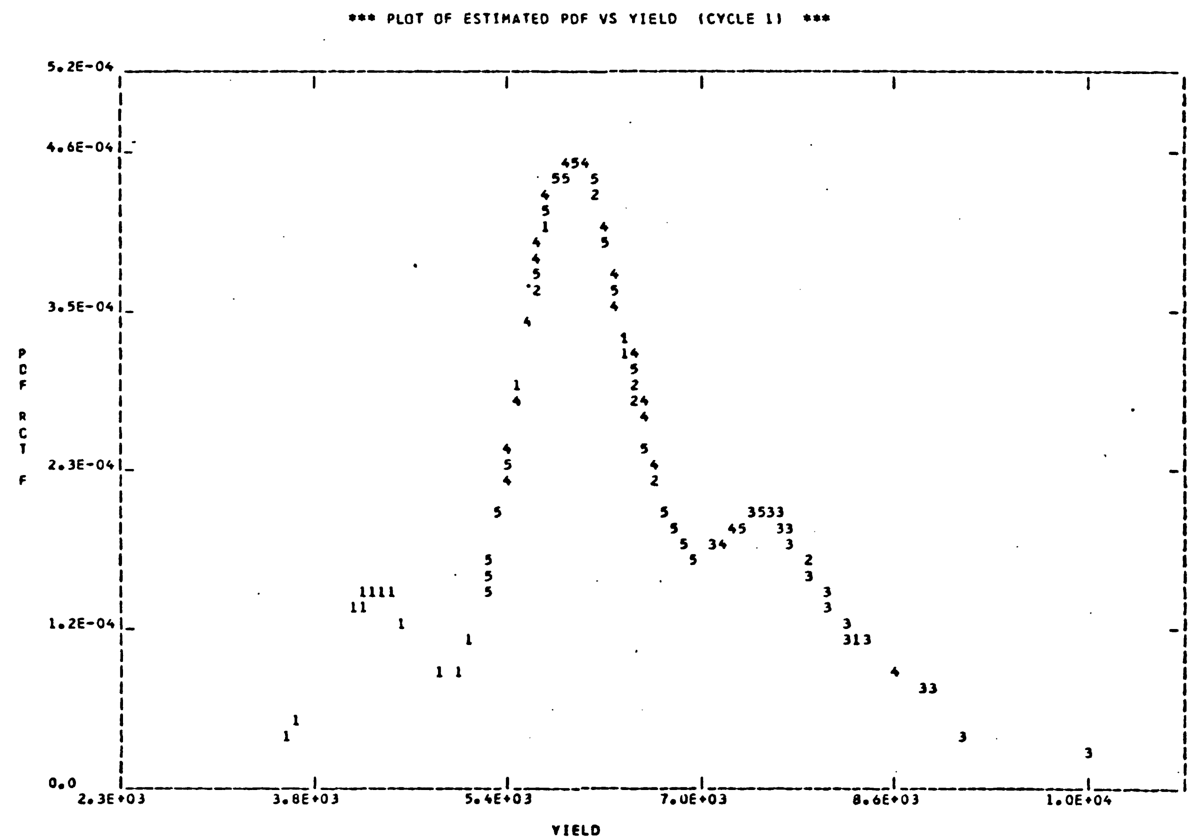
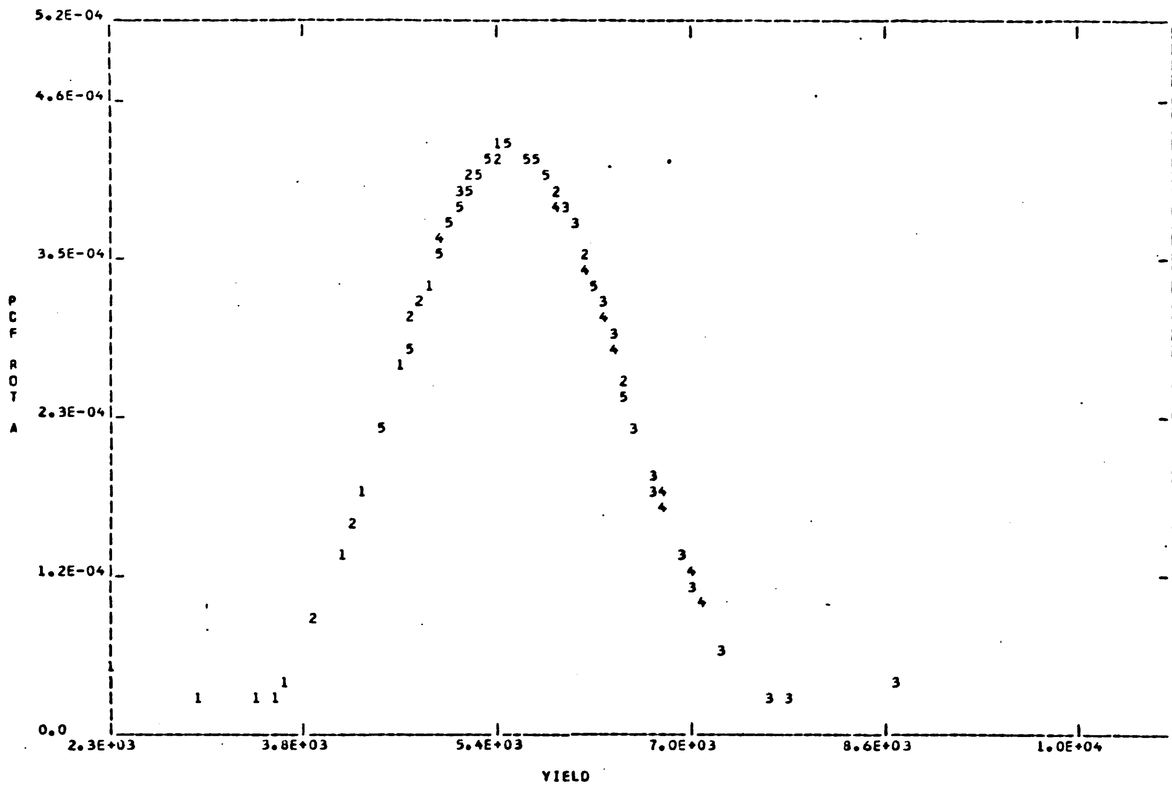

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***
PDF RCT F
YIELD
5.2E-04
4.6E-04
3.5E-04
2.3E-04
1.2E-04
0.0
2.3E+03
3.8E+03
5.4E+03
7.0E+03
8.6E+03
1.0E+04

Figure 2

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***




*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 1) ***

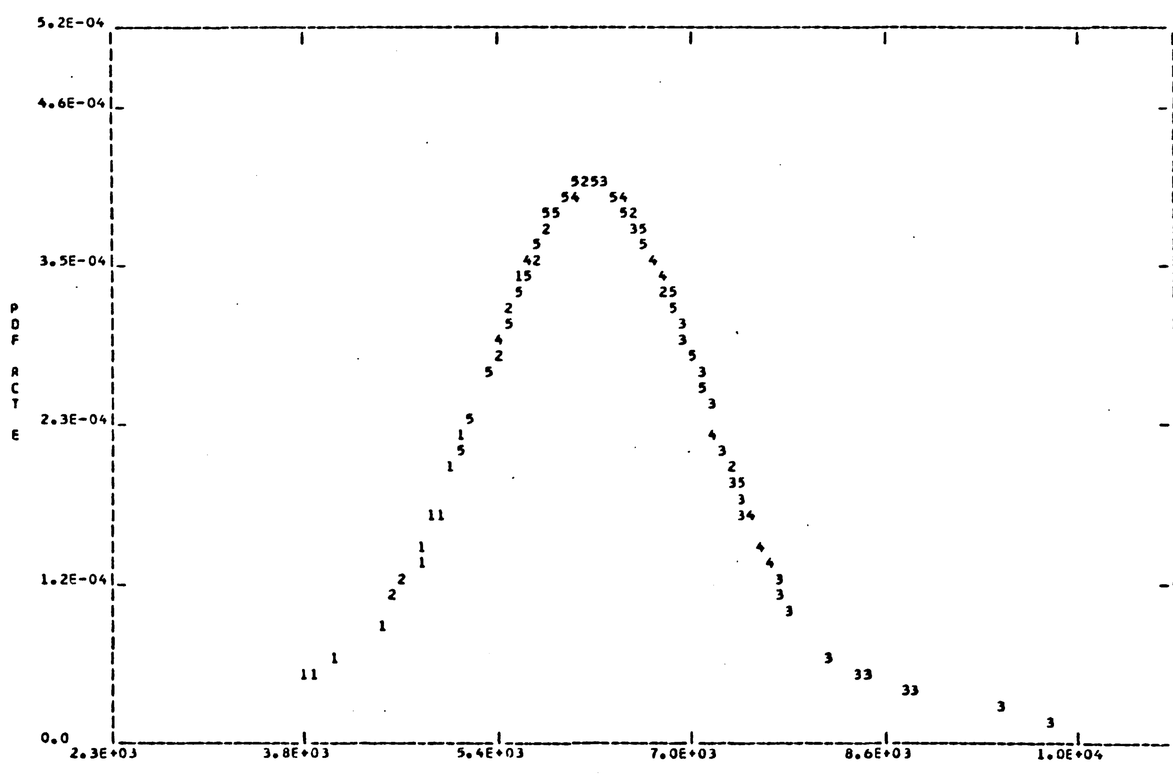



Figure 3

A: Continuous corn (plus manure),

B: Continuous corn (no manure),

C: Continuous corn after a C-C-O-A-A cycle during the previous five years,

E: C-O-W-A-A rotation,

F: C-O-Cl-W-A rotation,

G: C-O-A-A-A rotation,

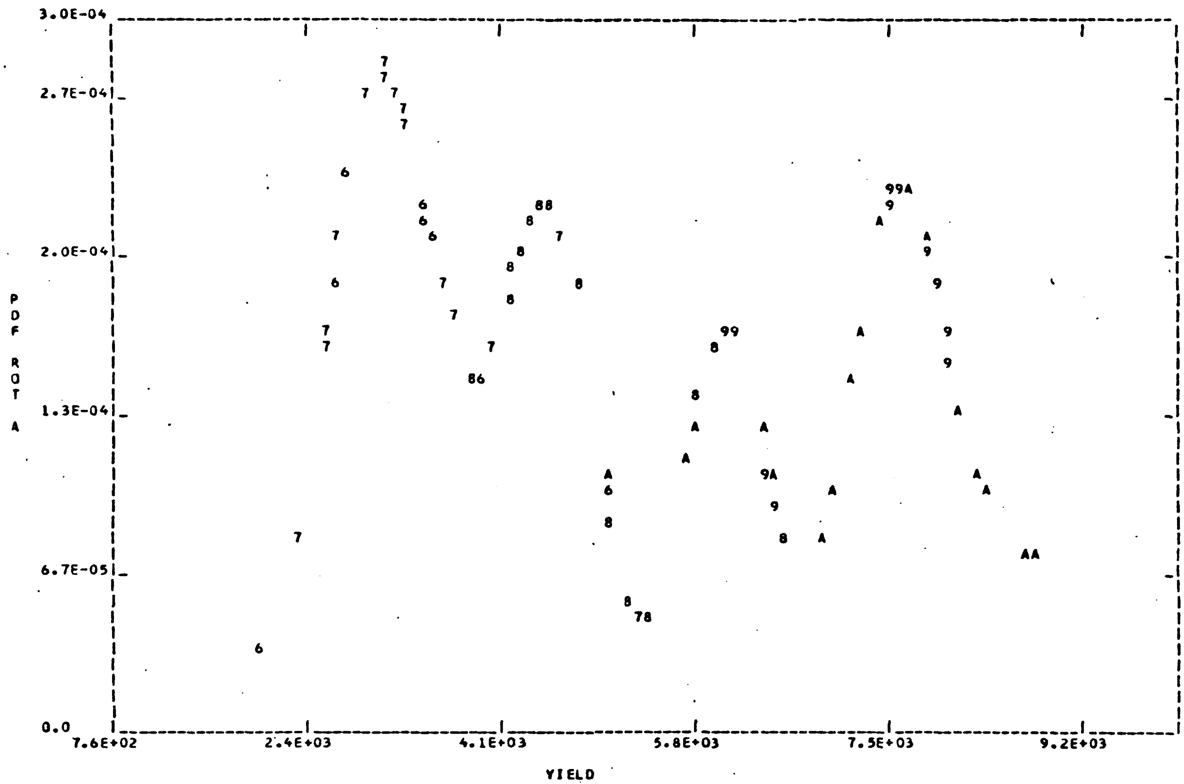
where C = corn, ) = oats, W = wheat, Cl = clover and A = alfalfa.

In terms of central tendency there is a noticeable shift towards higher yields in the plots for the five-year rotations as compared to the plots for the continuous rotations. Not only is there a shift in the location of the peak, but the whole area of the curve is shifted to the right indicating greater probabilities of high yields for the five-year rotation. The reaction of the five-year rotations to the one good climatic year (year 3) of cycle one is also quite different from that of the continuous rotations. Climatic details are given by Baldock (1976). The distribution of yields from these rotations form a secondary peak around a higher level of yield, indicating that the management policies of five-year rotations are benefiting yields in good climatic years. All of these differences are noticeable visually but are not large quantitatively.

The cycle two density plots (Figures 4-6) are dominated by the vastly different year effects. As before the five-year rotations seem to be producing slightly higher yields in the good years of the cycle. However, any differential response of the rotations to the year effects seems to be overshadowed by the year effects themselves.

Treatment effects generally are not visible on these plots except in two instances. Comparing the plots for rotation A and rotation B gives a comparison of manure (A) vs. no manure (B) in a continuous rotation. The comparison is in favor of A. It also seems that in cycle two of rotation A the best treatment has separated into a secondary peak in years 9 and 10. This is the only time that a treatment effect is strong enough to separate (identification not shown on the plots is needed to distinguish the treatments).

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***




*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***

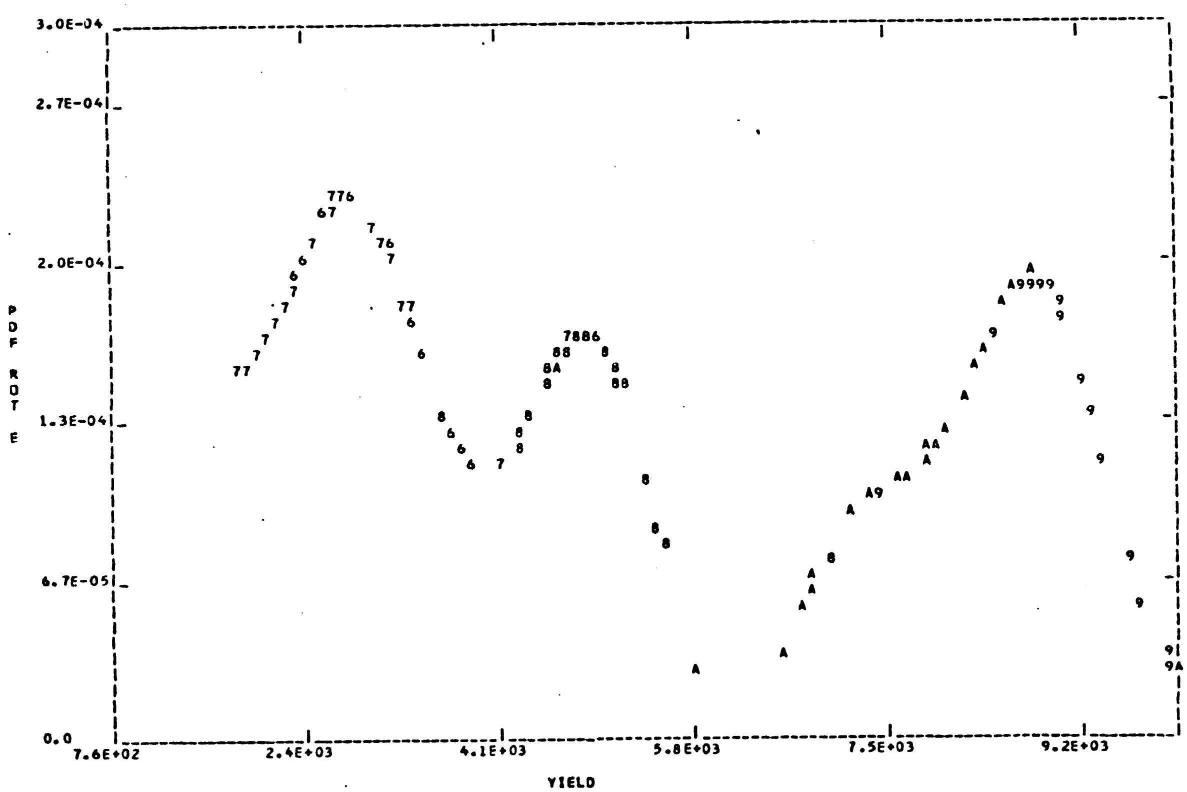
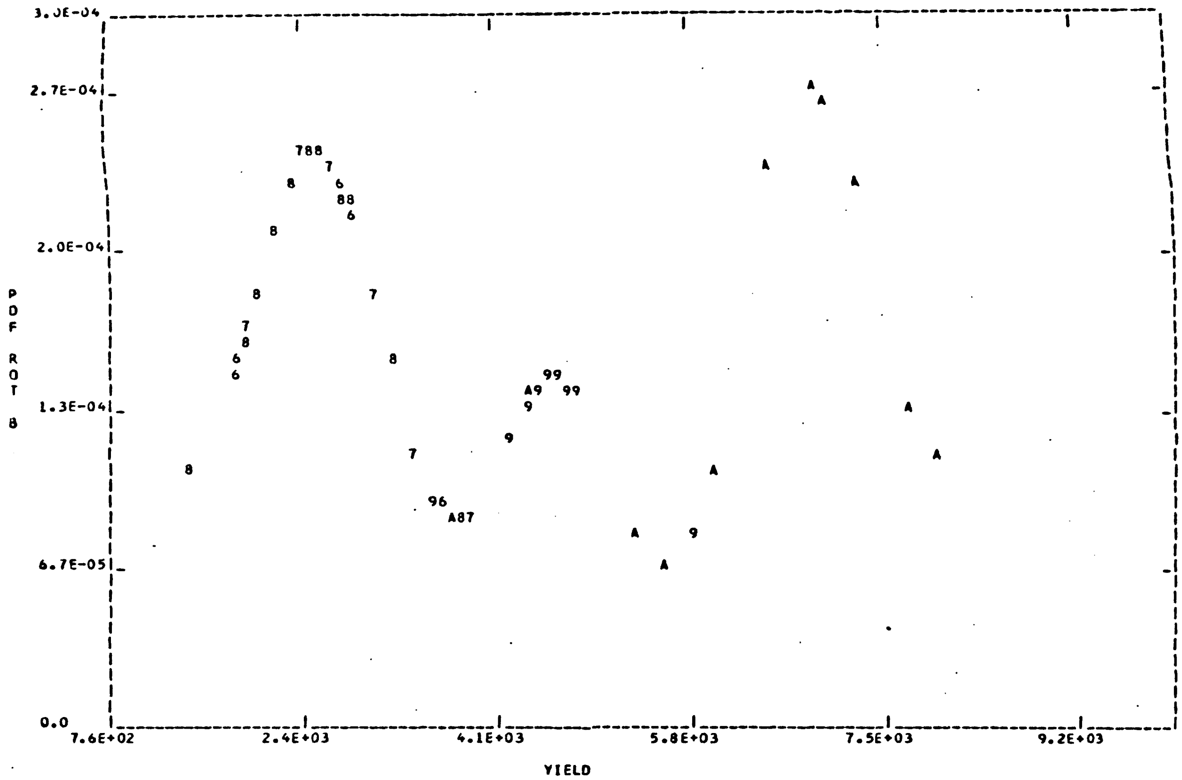



Figure 4

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***




*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***

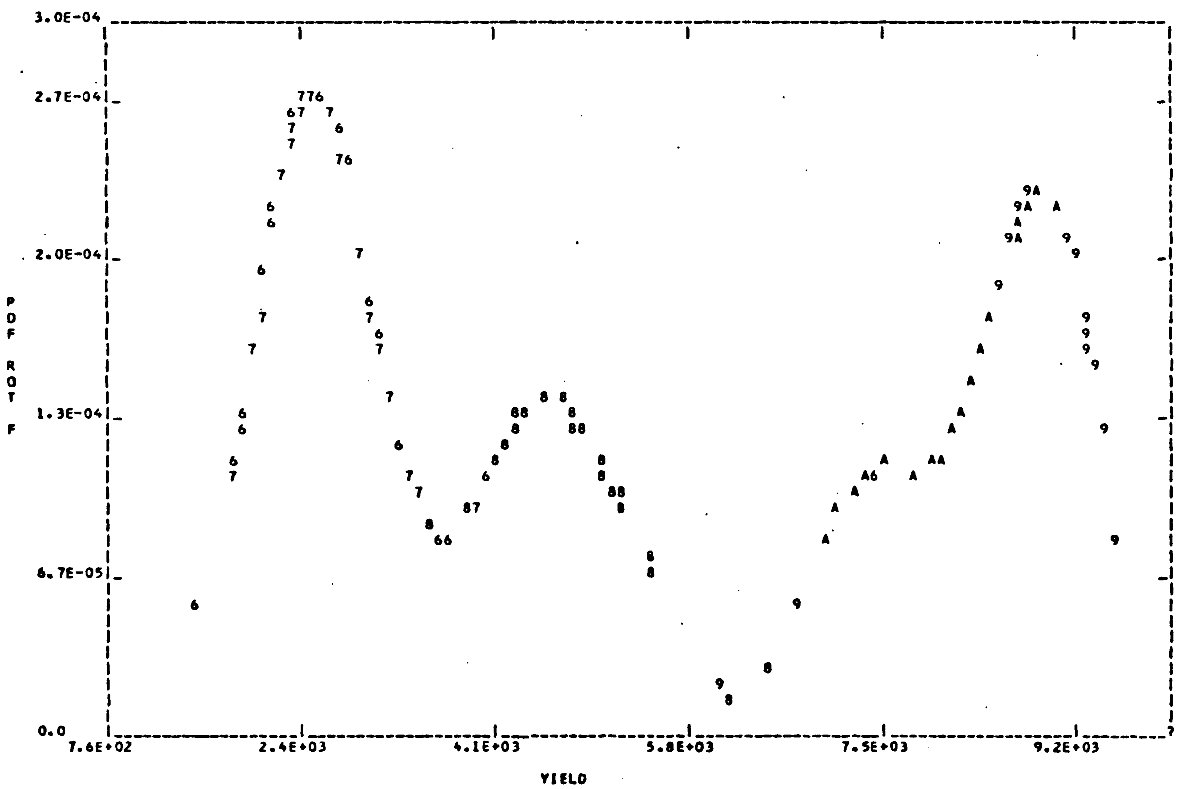
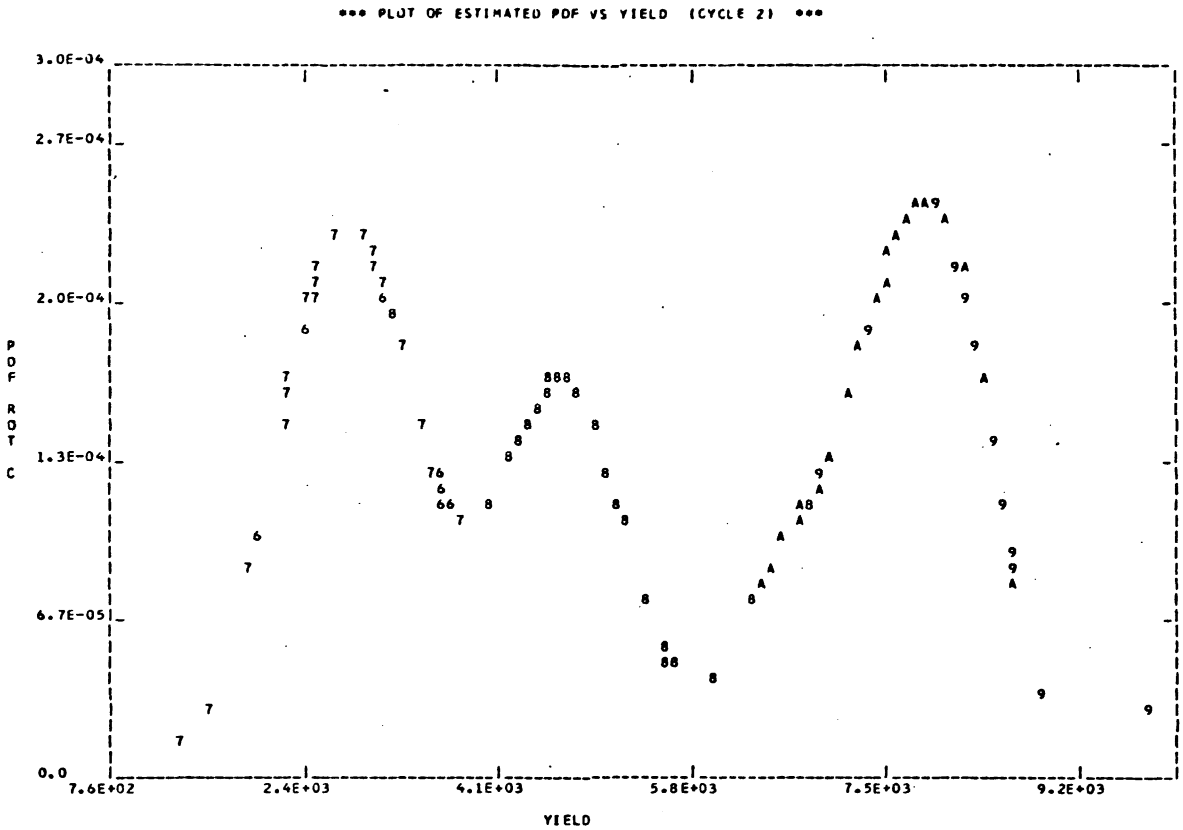



Figure 5


*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***
3.0E-04
2.7E-04
2.0E-04
1.3E-04
6.7E-05
0.0
PDF ROT C
7.6E+02
2.4E+03
4.1E+03
5.8E+03
7.5E+03
9.2E+03
YIELD

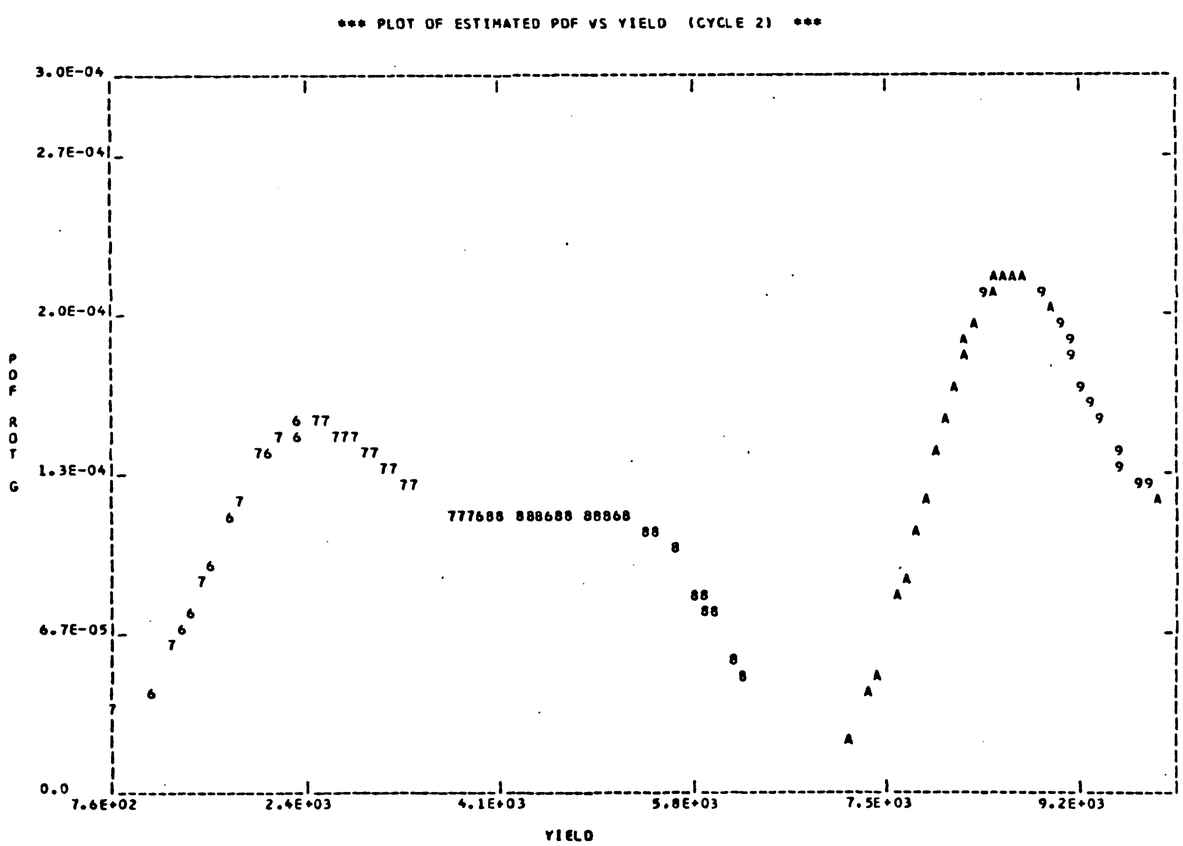

*** PLOT OF ESTIMATED PDF VS YIELD (CYCLE 2) ***
3.0E-04
2.7E-04
2.0E-04
1.3E-04
6.7E-05
0.0
PDF ROT G
7.6E+02
2.4E+03
4.1E+03
5.8E+03
7.5E+03
9.2E+03
YIELD

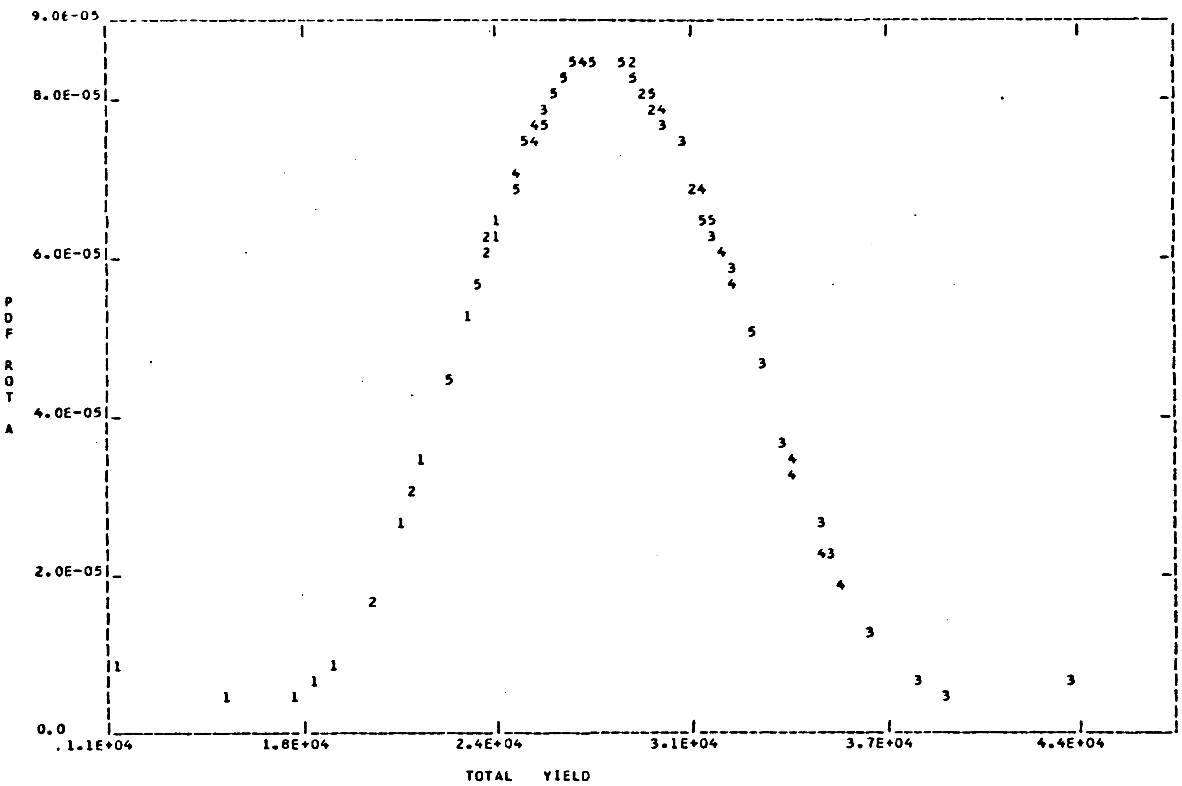
Figure 6

Density plots (Figures 7-9) are also available for the total dry matter yield from each of the rotations that contain corn. These plots generally give a much different comparison between the continuous and the five-year rotations. The only rotation which remains competitive with continuous corn cropping in terms of total dry matter yield is rotation G.

## 6.2. Cumulative Treatment Plots

The plots (Figures 10-13) of the cumulative yearly means for rotations A, B, CC and D1 (first-year corn from a C-C-O-A-A rotation) are self explanatory. With one exception, no treatment differences are evident from plots of the other rotations. The insert in the bottom right corner gives the numerical values of the means of the first five years, the means of the second five years, and the ten-year means for the corn yields from each rotation treatment combination. The motivation behind averaging over all previous years is to compress the effect of any single year or comparisons between the treatments. Moving from left to right across each graph (and hence looking at means of more and more data), the graphs should stabilize and give a consistent pattern of treatment effects. In fact, the year effects remain so large that the cumulative treatment means have not yet settled down but do seem to present a consistent pattern for treatments. The most dramatic comparisons are those between the manure and no-manure treatments. In general, the three applied nitrogen levels increase with the W, X = Y and Z treatments. With rotations C and D, the W and X treatments also have manure with rotation D at a higher rate. Table 2 summarizes the treatments and a complete description is given in Baldock (1976). One of the more interesting comparisons is between rotation A (continuous corn with manure) and rotation B. Superimposing the plot for rotation A over the plot for B, the moving average plots show the B means for treatments W and Z starting out considerably higher than the comparable

**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****




**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****

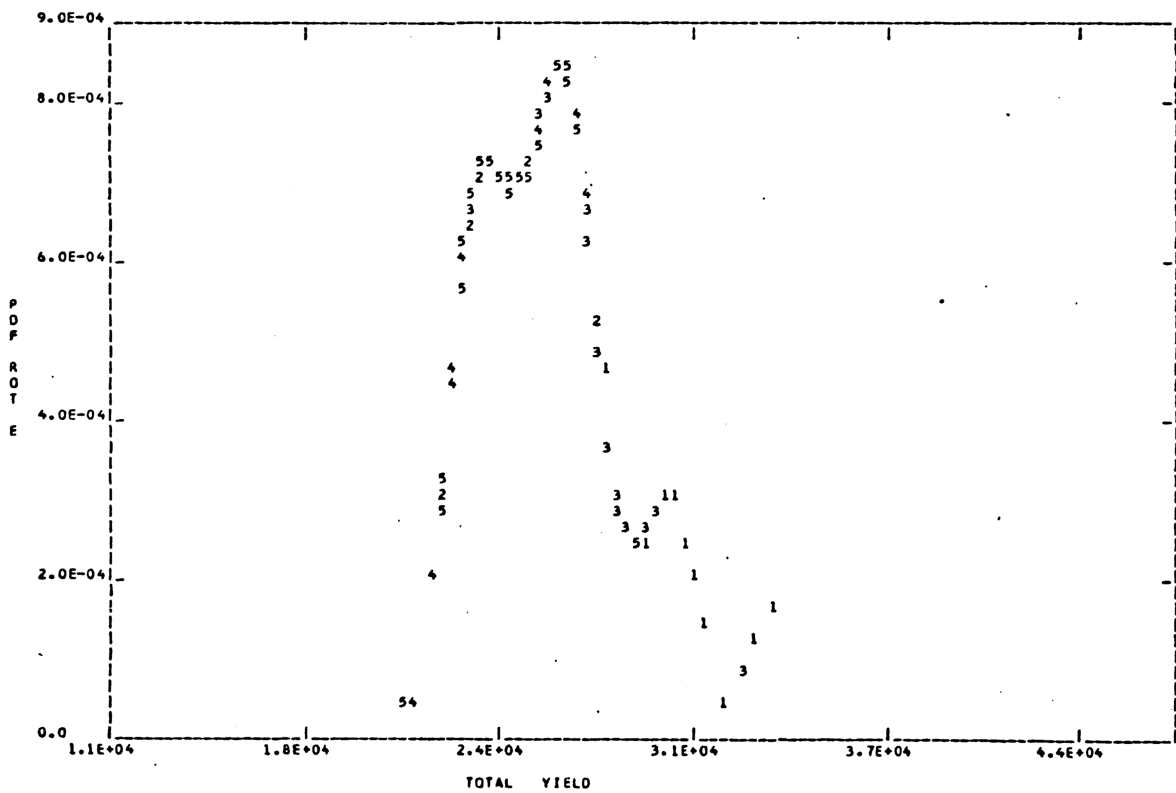
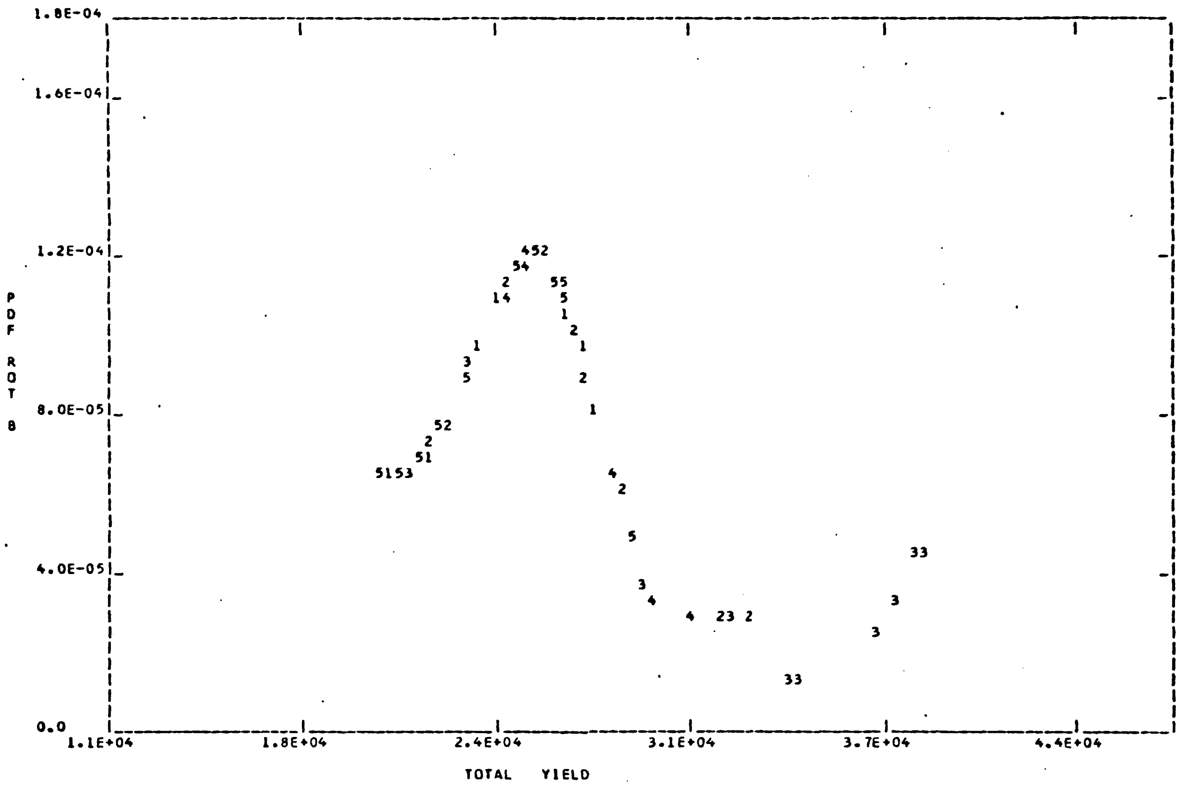



Figure 7

**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****




**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****

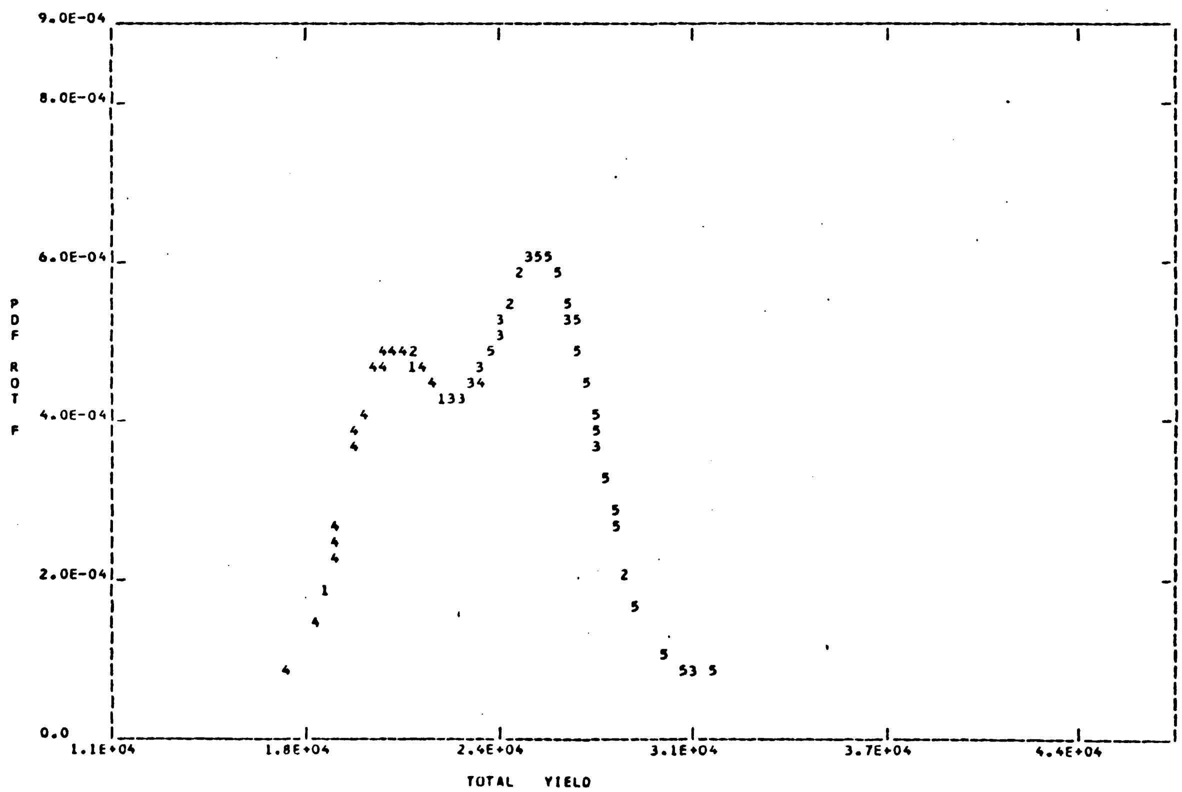
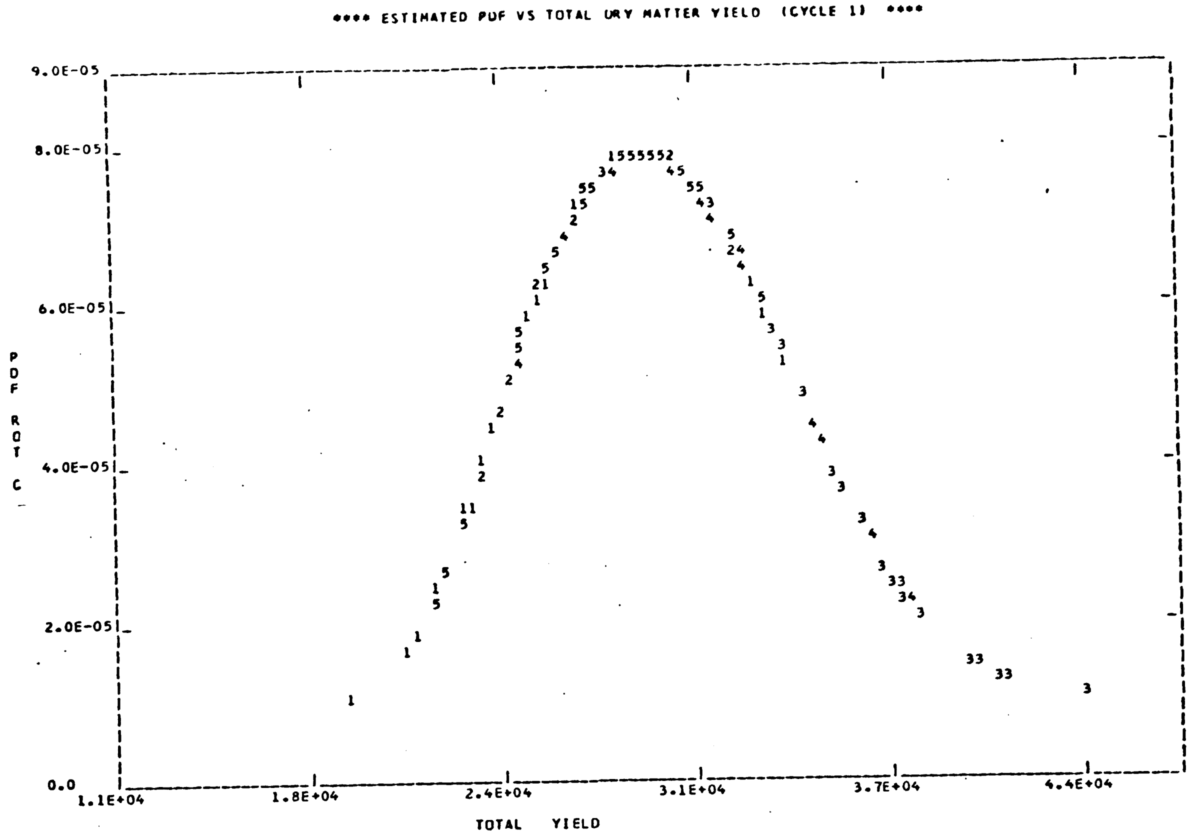



Figure 8


**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****
9.0E-05
8.0E-05
6.0E-05
4.0E-05
2.0E-05
0.0
PDF ROT C
1.1E+04
1.8E+04
2.4E+04
3.1E+04
3.7E+04
4.4E+04
TOTAL YIELD

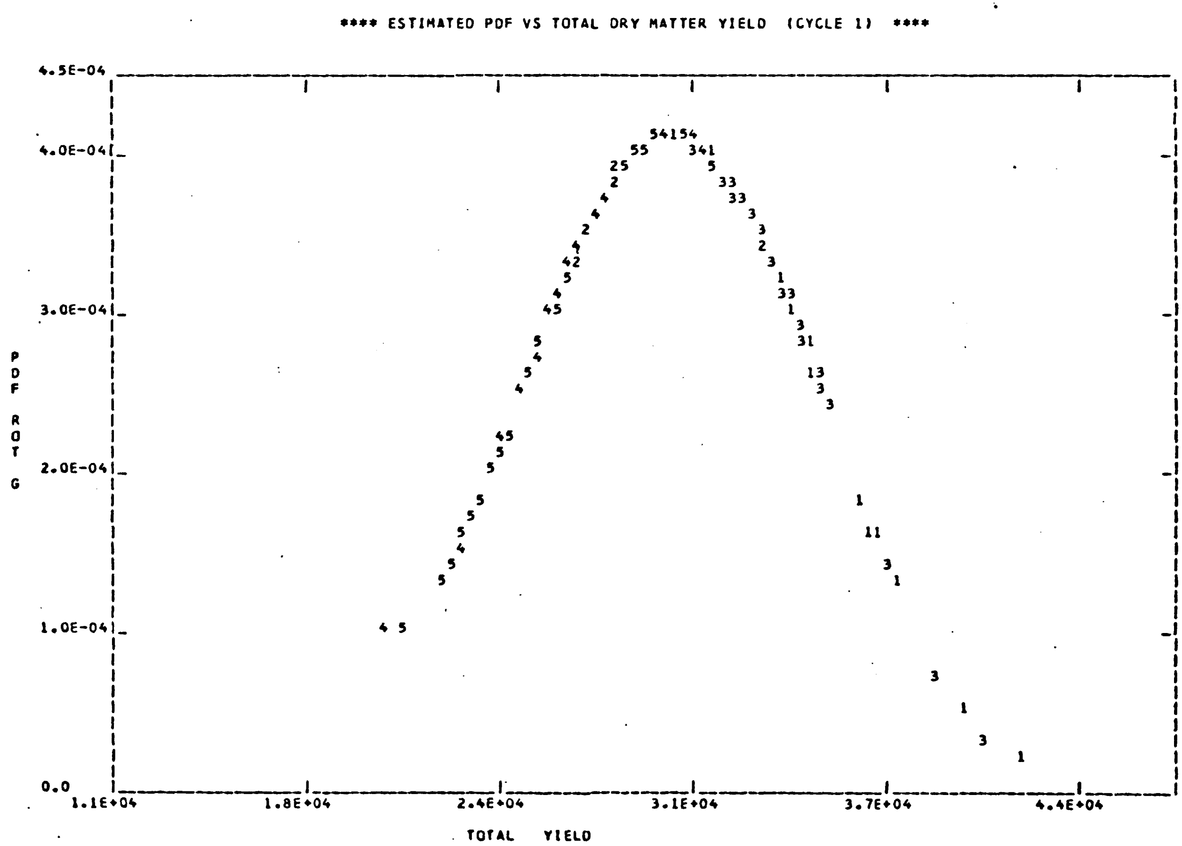

**** ESTIMATED PDF VS TOTAL DRY MATTER YIELD (CYCLE 1) ****
4.5E-04
4.0E-04
3.0E-04
2.0E-04
1.0E-04
0.0
PDF ROT G
1.1E+04
1.8E+04
2.4E+04
3.1E+04
3.7E+04
4.4E+04
TOTAL YIELD

Figure 9

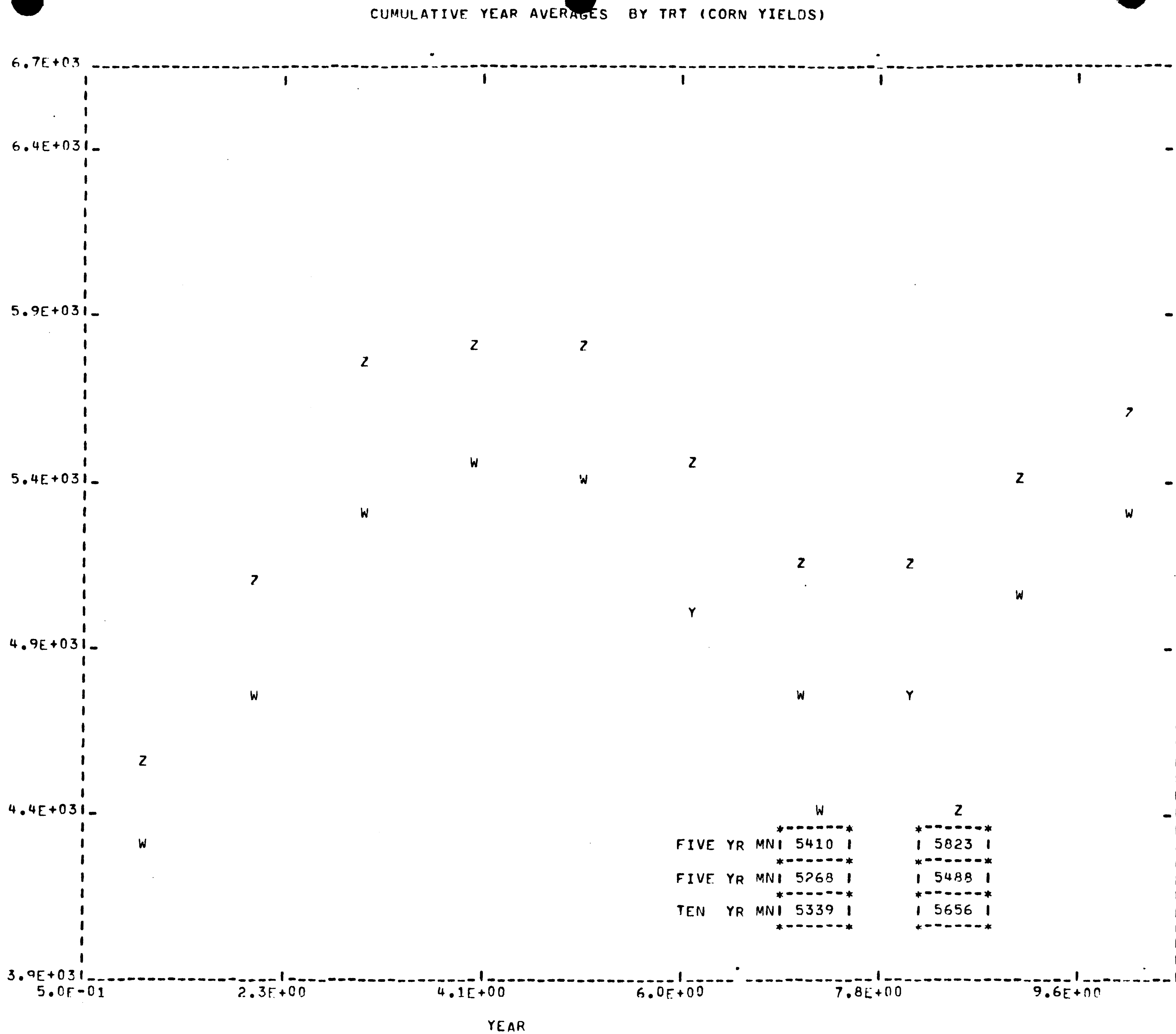

CUMULATIVE YEAR AVERAGES BY TRT (CORN YIELDS)
6.7E+03
6.4E+03
5.9E+03
5.4E+03
4.9E+03
4.4E+03
3.9E+03
ROT A
5.0E-01
2.3E+00
4.1E+00
6.0E+00
7.8E+00
9.6E+00
YEAR
W
Z
FIVE YR MN 5410 5823
FIVE YR MN 5268 5488
TEN YR MN 5339 5656

Figure 10

CUMULATIVE YEAR AVERAGES BY TRT (CORN YIELDS)

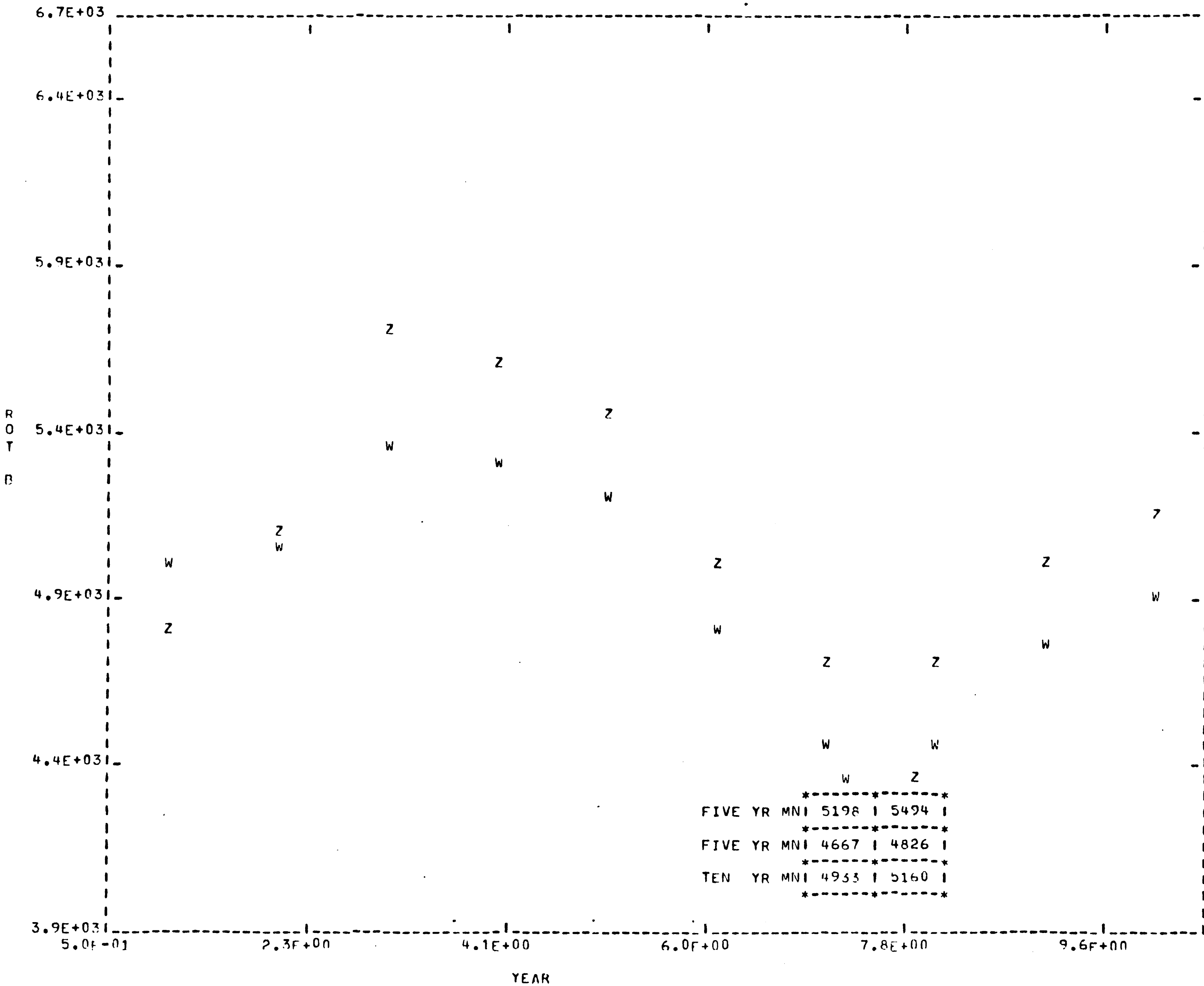
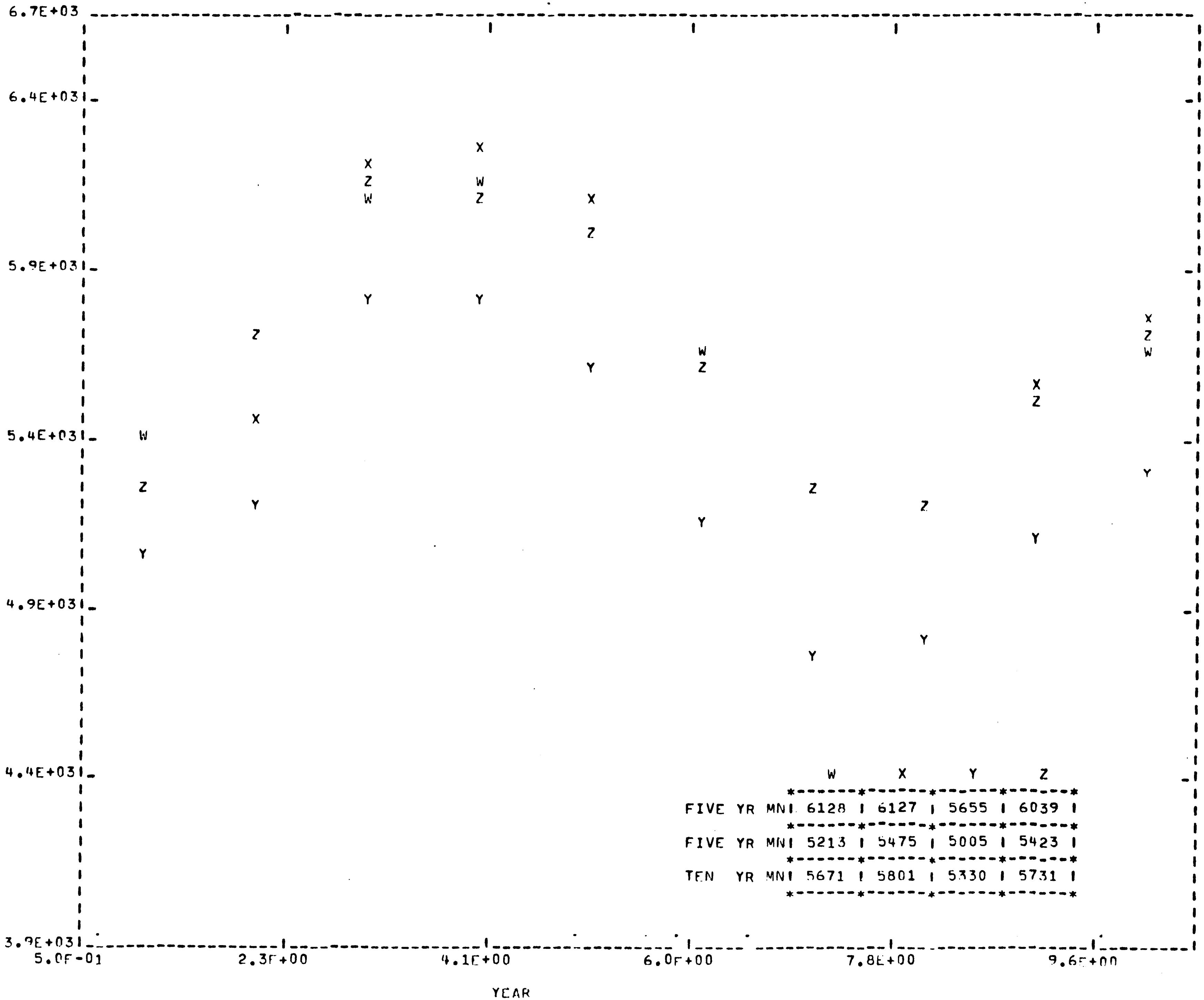



Figure 11

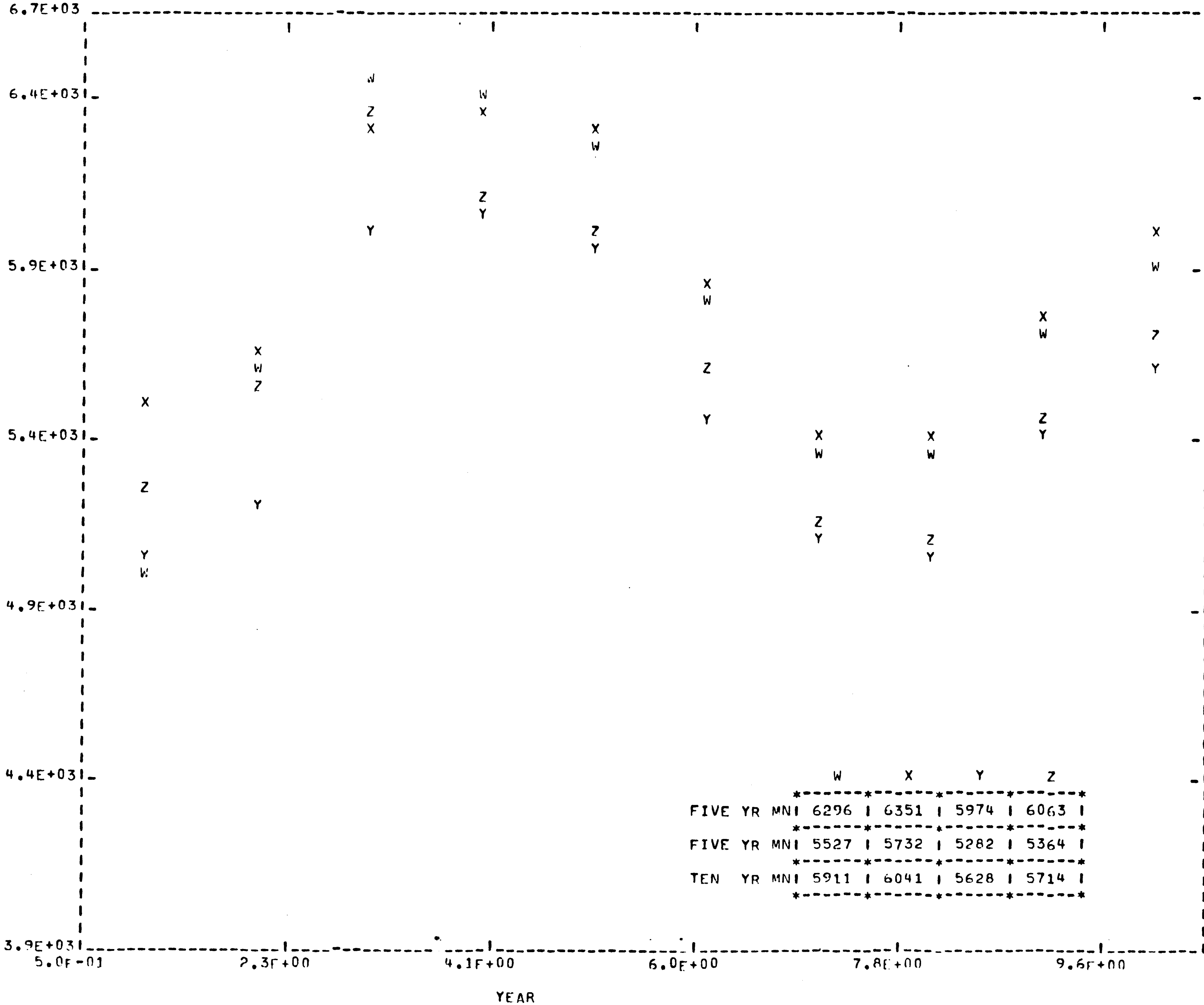
CUMULATIVE YEAR AVERAGES BY TRT (CORN YIELDS)




| | W | X | Y | Z |
|---|---|---|---|---|
| FIVE YR MN | 6128 | 6127 | 5655 | 6039 |
| FIVE YR MN | 5213 | 5475 | 5005 | 5423 |
| TEN YR MN | 5671 | 5801 | 5330 | 5731 |

Figure 12

CUMULATIVE YEAR AVERAGES BY TRT (CORN YIELDS)




| | W | X | Y | Z |
|---|---|---|---|---|
| FIVE YR MN | 6296 | 6351 | 5974 | 6063 |
| FIVE YR MN | 5527 | 5732 | 5282 | 5364 |
| TEN YR MN | 5911 | 6041 | 5628 | 5714 |

Figure 13

Table 2. Summary of Nitrogen and Manure Treatments

| Rotation | Treatments | | | |
|---|---|---|---|---|
| | W | X | Y | Z |
| A | normal N<br>manure | | | high N<br>manure |
| B | normal N<br>no manure | | | high N<br>no manure |
| C | low N<br>manure | normal N<br>manure | normal N<br>no manure | high N<br>no manure |
| D | low N<br>manure | normal N<br>manure | normal N<br>no manure | high N<br>no manure |

A treatment means. By 1961 the three-year treatment means are essentially comparable. However, as rotation A remains essentially stable in 1962 and 1963, the rotation B means decline and the difference between the rotation for both treatments tend to increase before stabilizing.

The ten-year averages show differences of $5339-4933 \cong 400$ at the lower level of applied N and $5656 - 5160 \cong 500$ at the higher level of applied N . The manure effect is also evident in comparing treatment X with Y for rotation C and D and comparing rotations C and D for treatments W and X .

## 7. Conclusions

(a) All the original raw field data have been edited and the final plot data are now available in a form accessible to other research and extension personnel interested in comparing rotations of particular interest to dairy farmers.

(b) Preliminary graphic analyses show the main features of the data including the comparison of certain five-year rotations with continuous corn rotations and the effects of fertility treatments, specifically a manure effect.

## 8. Acknowledgments

This research was supported by the New York State Experiment Station and covered part-time employment of R. H. Brown and computing expenses. Appreciation is expressed to Karen Rhodes for her assistance in the data management and to Jon Baldock, David Bouldin, L. W. McEachron, Robert Musgrave, Madison Wright, Paul Zwerman and the Agronomy Department for access to the original data and their support. One of the authors (F. B. Cady) has benefitted from earlier discussions concerning the analysis of rotation experiments with Chang-Sheng Shih and Wayne Fuller.

## 9. References

Baldock, J. O. (1976). What cropping system for New York dairy farms? An analysis of the Aurora rotation experiment. Ph.D. dissertation, Cornell University, Ithaca, New York.

Battese, G. F. and W. A. Fuller (1972). Determination of economic optimum from crop-rotation experiments. Biometrics 28, 781-792.

Battese, G. F., W. A. Fuller and W. D. Shrader (1972). Analysis of crop-rotation experiments, with application to the Iowa Carrington-Clyde rotation-fertility experiments. Iowa State University Research Bulletin 574, 579-595.

Björnsson, H. and F. B. Cady (1973). Correlated error structure in perennial crop experiments. Biometrics Unit Mimeo 469, Cornell University, Ithaca, New York.

Cady, F. B. and D. D. Mason (1964). Comparison of fertility treatments in a crop rotation experiment. Agronomy Journal 56, 476-479.

Cochran, W. G. (1939). Long-term agricultural experiments. Journal of the Royal Statistics Society Series B (Supplement) 6, 104-148.

Fuller, W. A. and G. F. Battese (1973). Transformations for estimation of linear models with nested-error structure. Journal of the American Statistical Association 68, 626-632.

Fuller, W. A. and F. B. Cady (1965). Estimation of asymptotic rotation and nitrogen effects. Agronomy Journal 57, 299-302.

Kronmal, R. A. and M. E. Tarter (1968). The estimation of probability densities and cumulatives by Fourier series methods. Journal of the American Statistical Association 63, 925-952.

Patterson, H. D. (1959). The analysis of a non-replicated experiment involving a single 4-course rotation of crops. Biometrics 15, 30-59.

Patterson, H. D. (1964). Theory of cyclic experiments. Journal of the Royal Statistical Society, Series B 26, 1-36 (Discussion pp. 36-43).

Patterson, H. D. and B. I. Lowe (1970). The errors of long-term experiments. Journal of Agricultural Science 74, 53-60.

Shih, C. S. (1966). Interval estimation for the exponential model and the analysis of rotation experiments. Ph.D. thesis, Iowa State University, 126 pp. (Mic. 66-10440, University Microfilms, Ann Arbor, Michigan)(Diss. Abstract 27 (4B), 1318).

Shrader, W. D., W. A. Fuller and F. B. Cady (1966). Estimation of a common nitrogen response function for corn in different crop rotations. Agronomy Journal 28, 397-401.

Tarter, M. E. and R. A. Kronmal (1976). An introduction to the implementation and theory of nonparametric density estimation. American Statistician 30, 105-112.

Tarter, M. E., R. L. Holcomb and R. A. Kronmal (1967). A description of new computer methods for estimating the population density. In Proceedings, Association for Computing Machinery 22, 511-519. Thompson Book Company, Washington, D. C.

Yates, F. (1954). The analysis of experiments containing different crop rotations. Biometrics 10, 324-346.